

**KOMPARASI METODE *SUPPORT VECTOR MACHINE (SVM)* DENGAN
BIDIRECTIONAL ENCODER REPRESENTATION TRANSFORMERS
(*BERT*) UNTUK KLASIFIKASI SENTIMEN KOMENTAR VIDEO
YOUTUBE MATA NAJWA**



OLEH

ANISZA ASMARA KELANA

NIM 1915101030

PROGRAM STUDI ILMU KOMPUTER

JURUSAN TEKNIK INFORMATIKA

FAKULTAS TEKNIK DAN KEJURUAN

UNIVERSITAS PENDIDIKAN GANESHA


SINGARAJA

2025



**KOMPARASI METODE *SUPPORT VECTOR MACHINE (SVM)*
DENGAN *BIDIRECTIONAL ENCODER REPRESENTATION
TRANSFORMERS(BERT)* UNTUK KLASIFIKASI
SENTIMEN KOMENTAR VIDEO YOUTUBE
MATA NAJWA**

SKRIPSI



**Diajukan Kepada
Universitas Pendidikan Ganesha
Untuk memenuhi salah satu persyaratan
Dalam Menyelesaikan Program Sarjana
Program Studi Ilmu Komputer**

**Oleh
ANISZA ASMARA KELANA**

NIM 1915101030

**PROGRAM STUDI ILMU KOMPUTER
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA
SINGARAJA
2025**

SKRIPSI

**DIAJUKAN UNTUK MELENGKAPI TUGAS-TUGAS
DAN MEMENUHI SYARAT-SYARAT UNTUK
MENCAPAI GELAR SARJANA KOMPUTER**


Menyetujui,

Pembimbing I,



Ni Putu Novita Puspa Dewi, S.Kom., M.Cs.
NIP. 199410032020122015

Pembimbing II,



Dr. Putu Hendra Suputra, S.Kom., M.Cs.
NIP. 198212222006041001

Skripsi oleh **Anisza Asmara Kelana** ini telah dipertahankan di depan dewan penguji pada tanggal : 07 Februari 2025

Dewan Penguji,



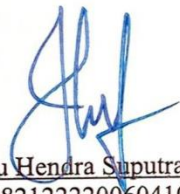
I Ketut Resika Arthana, S.T., M.Kom.
NIP. 198412012012121002

(Ketua)



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 198910262019031004

(Anggota)



Dr. Putu Hendra Suputra, S.Kom., M.Cs.
NIP. 198212222006041001

(Anggota)



Ni Putu Novita Puspa Dewi, S.Kom., M.Cs.
NIP. 199410032020122015

(Anggota)

Diterima oleh Panitia Ujian Fakultas Teknik dan Kejuruan
Universitas Pendidikan Ganesha
guna memenuhi syarat-syarat untuk mencapai gelar sarjana komputer

Pada :

Hari : *Kamis*

Tanggal : *13 FEB 2025*



Mengetahui,

Ketua Ujian,

Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.
NIP. 198211112008121001

Sekretaris Ujian,

I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 198910262019031004

Mengesahkan

Dekan Fakultas Teknik dan Kejuruan

Prof. Dr. Kadek Rihendra Dantes, S.T., M.T.
NIP. 197912012006041001

PERNYATAAN

Dengan ini saya menyatakan bahwa karya tulis yang berjudul “Komparasi Metode *Support Vector Machine* (SVM) Dengan *Bidirectional Encoder Representation Transformers* (BERT) Untuk Klasifikasi Sentimen Komentar Video Youtube Mata Najwa” beserta seluruh isinya adalah benar-benar karya sendiri dan saya tidak melakukan penjiplakan dan pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung resiko/sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran atas etika keilmuan dalam karya saya ini atau ada klaim terhadap keaslian karya saya ini.

Singaraja, 13 Februari 2025.



Anisza Asmara Kelana
NIM. 1915101030

“Berusaha Menyelesaikan Apa Yang Sudah Dimulai”



PRAKATA

Puji syukur penyusun panjatkan ke hadapan Tuhan Yang Maha Esa karena berkat rahmat-Nya-lah, penulis dapat menyelesaikan skripsi yang berjudul **“Komparasi Metode *Support Vector Machine* (SVM) Dengan *Bidirectional Encoder Representation Transformers* (BERT) Untuk Klasifikasi Sentimen Komentar Video Youtube Mata Najwa”**.

Skripsi ini diajukan untuk memenuhi syarat kelulusan mata kuliah skripsi di Fakultas Teknik dan Kejuruan Undiksha Singaraja. Dalam menyelesaikan skripsi ini, penulis banyak mendapat bantuan baik berupa moral maupun material dari berbagai pihak. Untuk itu, dalam kesempatan ini, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. I Wayan Lasmawan M.Pd., selaku Rektor Universitas Pendidikan Ganesha atas motivasi dan fasilitas yang diberikan sehingga penulis dapat menyelesaikan studi sesuai dengan rencana.
2. Bapak Prof. Dr. Gede Rasben Dantes, S.T, M.T.I. selaku Wakil Rektor 1 Bidang Akademik dan Kerjasama atas motivasi yang diberikan dalam penyelesaian skripsi ini.
3. Bapak Dr. Kadek Rihendra Dantes, S.T., M.T. selaku Dekan Fakultas Teknik dan Kejuruan atas motivasi dan fasilitas yang diberikan sehingga penulis dapat menyelesaikan studi sesuai dengan rencana.
4. Bapak Prof. Dr. Komang Setemen, S.Si., M.T. selaku Kepala Biro Bidang Perencanaan dan Kerjasama atas motivasi yang diberikan dalam penyelesaian skripsi ini.
5. Bapak Dr. Putu Hendra Suputra, S.Kom., M.Cs., selaku Ketua Jurusan Teknik Informatika sekaligus Pembimbing II yang telah memberikan arahan, petunjuk skripsi dan motivasi penulis dalam penyelesaian skripsi ini.
6. Bapak I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs. selaku Ketua Prodi Ilmu Komputer sekaligus Penguji II atas motivasi dan bimbingan yang diberikan dalam penyelesaian skripsi ini.

7. Bapak I Ketut Resika Arthana, S.T., M.Kom. selaku Penguji I atas motivasi dan bimbingan yang diberikan dalam penyelesaian skripsi ini.
8. Ibu Ni Putu Novita Puspa Dewi, S.Kom., M.Cs. selaku Pembimbing I yang telah memberikan bimbingan, arahan, petunjuk skripsi dan motivasi penulis dalam penyelesaian skripsi ini.
9. Seluruh Dosen Jurusan Teknik Informatika/Program Studi Ilmu Komputer yang telah berbagi ilmu dan pengalamannya selama saya kuliah semoga semua yang peneliti pelajari di kampus dapat bermanfaat bagi masyarakat luas.
10. Untuk orang tua saya Kustono dan Yatini, serta ke dua kakak saya Ariya Asmara Kelana dan Gana Rendra Winarti yang senantiasa memberikan dukungan dan motivasi selama menempuh pendidikan di Universitas Pendidikan Ganesha.
11. Kepada Anotator dan Validator membatu kegiatan peneliti selama melakukan penulisan dan penyelesaian skripsi ini.
12. Kepada sahabat saya zahra, meirlianah, enung, dan regina yang telah memberikan semangat, dan motivasi kepada peneliti selama melakukan penulisan dan penyelesaian skripsi ini.
13. Untuk seluruh teman-teman Program Studi Ilmu Komputer Angkatan 2019 yang telah memberikan berbagai pengalaman dan kebersamaan selama peneliti menempuh pendidikan di Universitas Pendidikan Ganesha.
14. Semua pihak yang telah membantu penulis dari saudara, sahabat, dan orang terdekat lainnya yang telah menuntun, memberikan dukungan, masukan, saran dan hiburan yang tidak dapat saya sebutkan satu persatu. Saya ucapkan terima kasih sebanyak-banyaknya.
15. Peneliti menyadari sepenuhnya bahwa apa yang disajikan dalam skripsi ini masih perlu diperbaiki mengingat keterbatasan kemampuan peneliti. Untuk itu demi kesempurnaan skripsi, peneliti mengharapkan segala kritik dan saran yang bersifat membangun dari berbagai pihak. Peneliti berharap skripsi ini dapat bermanfaat dan berharga bagi kita semua, khususnya untuk mengembangkan dunia pendidikan.

Singaraja,

Peneliti



DAFTAR ISI

SKRIPSI	ii
PERNYATAAN	vi
PRAKATA	viii
ABSTRAK	xi
ABSTRACT	xii
DAFTAR ISI	xiii
DARTAR TABEL	xvi
DAFTAR GAMBAR	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	4
BAB II KAJIAN TEORI	5
2.1 Penelitian Terkait	5
2.3 Analisis Sentimen	8
2.4 Support Vector Machine	8
2.5 Bidirectional Encoder Representation From Transformers	10
2.6 Encoder	11
2.7 Decoder	11
2.8 Scraping Data	16
2.9 Preprocessing	17

2.12 Case Folding.....	17
2.13 Cleansing	18
2.14 Tokenisasi.....	18
2.15 Normalisasi	19
2.16 Stopword Removal	19
2.17 Stemming.....	20
2.18 Word2vec	21
2.19 Confusion Matrix	22
2.20 Visualisasi Data	25
BAB III METODOLOGI PENELITIAN	26
3.1 Objek Penelitian	26
3.2 Tahapan Penelitian	26
3.3 Scraping Data Komentar Youtube	29
3.4 Pelabelan Data	30
3.5 Text Preprocessing	31
3.6 Word2vec Text Embedding.....	33
3.7 Pembagian Dataset.....	34
3.8 Klasifikasi Sentimen	35
3.9 Pengujian Model	35
3.10 Evaluasi performa	36
3.11 Visualisasi World Cloud	38
BAB IV PEMBAHASAN.....	39
4.1 Dataset.....	39

4.2	Pelabelan Data	39
4.3	Preprocessing	41
4.3.1	Case Folding.....	41
4.3.2	Cleaning	43
4.3.3	Tokenize	45
4.3.4	Stopword Removal.....	47
4.3.5	Stemming	49
BAB V	PENUTUP	72
DAFTAR PUSTAKA	75



DARTAR TABEL

Tabel 2. 1 Contoh Penerapan Case Folding	17
Tabel 2. 2 Contoh Penerapan Cleansing	18
Tabel 2. 3 Contoh Penerapan Tokenisasi	19
Tabel 2. 4 Contoh Penerapan Normalisasi	19
Tabel 2. 5 Contoh Penerapan Stopword Removal	20
Tabel 2. 6 Contoh Penerapan Stemming	21
Tabel 2. 7 Confusion Matrix	22
Tabel 3. 1 Contoh hasil Pelabelan Data	30
Tabel 3. 2 Contoh hasil Case Folding	31
Tabel 3. 3 Contoh hasil Cleaning	32
Tabel 3. 4 Contoh hasil Tokenize	32
Tabel 3. 5 Contoh hasil Stopword Removal	33
Tabel 3. 6 Contoh hasil Stemming	33
Tabel 4. 1 Hasil Data yang telah dilabeli	40
Tabel 4. 2 Jumlah data yang telah diberi label	40
Tabel 4. 3 Perbedaan data sesudah dan sebelum Case folding	42
Tabel 4. 4 Perbedaan data sesudah dan sebelum Cleaning	44
Tabel 4. 5 Perbedaan data sesudah dan sebelum Tokenize	46
Tabel 4. 6 Perbedaan data sesudah dan sebelum Stopword Removal	47
Tabel 4. 7 Perbedaan sesudah dan sebelum Stemming	50
Tabel 4. 8 Hasil Training dan Testing	51
Tabel 4. 9 Hasil Training dan Testing SVM Dengan Metode Down Sampling	56
Tabel 4. 10 Tokenize text menggunakan BERT	62
Tabel 4. 11 Hasil Training dan Testing BERT Dengan Semua Data	62
Tabel 4. 12 Hasil Training dan Testing BERT Dengan Metode Down Sampling	65

DAFTAR GAMBAR

Gambar 2. 1 Contoh Pemisahan Kelas Positif dan Negatif Hyperplane SVM	10
Gambar 2. 2 Encoder dan Decoder	12
Gambar 2. 3 Proses pada self-attention Sumber : Alammr (2018)	13
Gambar 2. 4 Proses pada Encoder Sumber: Alammr (2018)	14
Gambar 2. 5 Perbedaan ukuran pada BERT	15
Gambar 2. 6 Arsitektur BERT	16
Gambar 3. 1 Tahapan Penelitian Metode SVM	27
Gambar 3. 2 Tahapan Penelitian Metode BERT	28
Gambar 3. 3 Tahapan Scraping Data	29
Gambar 3. 4 Hasil Scraping Data.....	30
Gambar 3. 5 Diagram ROC - AOC.....	38
Gambar 4. 1 Persentasi hasil labeling data	41
Gambar 4. 2 Implementasi source code untuk case folding.....	42
Gambar 4. 3 Implementasi source code untuk Cleaning.....	43
Gambar 4. 4 Implementasi source code untuk Tokenize	45
Gambar 4. 5 Implementasi source code untuk Stopword Removal	47
Gambar 4. 6 Implementasi source code untuk Stemming.....	49
Gambar 4. 7 Hasil Accuracy Training SVM Dengan Semua Data.....	51
Gambar 4. 8 Confusion Matrix Training SVM Dengan Semua Data	53
Gambar 4. 9 Hasil Accuracy testing SVM Dengan Semua Data	54
Gambar 4. 10 Confusion Matrix Testing SVM Dengan Semua Data	55
Gambar 4. 11 Hasil Accuracy Training SVM Dengan Metode Down Sampling...56	
Gambar 4. 12 Confusion matrix Training SVM Dengan Metode Down Sampling57	
Gambar 4. 13 Hasil Accuracy Testing SVM Dengan Metode Down Sampling.....58	
Gambar 4. 14 Confusion Matrix Testing SVM Dengan Metode Down Sampling.58	
Gambar 4. 15 Visualisasi Grafik Kinerja Model Dengan Metode Down Sampling	60

Gambar 4. 16 Visualisasi Grafik ROC Dengan Metode Down Sampling.....	60
Gambar 4. 17 Hasil Accuracy Training BERT Dengan Semua Data	62
Gambar 4. 18 Confusion matrix training BERT Dengan Semua Data	64
Gambar 4. 19 Hasil Accuracy Testing BERT Dengan Semua Data.....	65
Gambar 4. 20 Confusion matrix testing BERT Dengan Semua Data.....	65
Gambar 4. 21 Hasil Accuracy Training BERT Dengan Metode Down Sampling .	66
Gambar 4. 22 Confusion Matrix Ttraining BERT Dengan Metode Down Sampling	66
Gambar 4. 23 Hasil Accuracy Testing BERT Dengan Metode Down Sampling ...	67
Gambar 4. 24 Confusion Matrix Testing BERT Dengan Metode Down Sampling	67
Gambar 4. 25 Hasil Trainig Loss BERT Dengan Metode Down Sampling	68
Gambar 4. 26 Hasil Akurasi ROC BERT Dengan Metode Down Sampling	69
Gambar 4. 27 Hasil Visualisasi World Cloud SVM	71
Gambar 4. 28 Hasil Visualisasi World Cloud BERT Tanpa Menggunakan Stopword	71



